Spatial gaps in global biodiversity information and the role of citizen science

Tatsuya Amano, James Lamming and William J. Sutherland

Conservation Science Group, Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK

Corresponding author: Amano, T. (amatatsu830@gmail.com).

Keywords: biodiversity data, conservation science, Global Biodiversity Information Facility (GBIF), information bias, knowledge gap

Abstract

Due to a range of constraints, the availability of biodiversity-related information varies considerably over space, time, taxa and types of data, thus causing gaps in knowledge. Despite growing awareness of this issue among scientists, it is still poorly known how, and whether, scientific efforts have contributed towards overcoming these information gaps. Focusing on spatial gaps in global biodiversity data, we show that accumulation rates of non-bird species occurrence records stored in the Global Biodiversity Information Facility have not improved, and have even potentially declined, over the past three decades in data-poor, often biodiversity-rich, regions. Meanwhile, one citizen science project, eBird, has been making a considerable contribution to the collection and sharing of bird records, even in data-poorest countries and is accelerating the accumulation of bird records globally. We discuss the potentials and limitations of citizen science projects for tackling gaps in biodiversity information, particularly from the perspective of biodiversity conservation.

Information gaps in conservation

With continually advancing research in the studies of ecology and conservation, we have accumulated considerable knowledge of species inhabiting this planet. Nevertheless, the availability of scientific information is affected by a range of factors, such as socioeconomic status, history, culture, geography and scientific interests, and thus varies greatly over space, time, taxa and types of information, creating gaps in biodiversity information. For example, the unequal distribution of biodiversity data across the globe, particularly the lack of information in biodiversity-rich regions, has repeatedly been reported since the 1980s (Western et al. 1989, Amano and Sutherland 2013, Pimm et al. 2014). Similarly, the availability of scientific data greatly varies over time (e.g., Gardner et al. 2014); information gaps can also be found in the coverage of taxa and ecosystems. In the assessment of species extinction risk by the International Union for Conservation of Nature (IUCN), only 0.6% of

10,425 bird species, but 46% of 1,084 cartilaginous fish species, are classified as Data Deficient (IUCN 2015). Some biomes, such as tropical deciduous woodlands and deserts, are typically underrepresented in ecological studies (Martin et al. 2012), as are marine systems in the IUCN assessments (Webb and Mindel 2015). Moreover, there are inevitable gaps in the types of available information. Long-term, broad-scale, standardised monitoring data, which are useful for deriving robust scientific inferences, are not common (Isaac et al. 2014), while less structured opportunistic data are now rapidly being accumulated thanks to the development of global databases, such as the Global Biodiversity Information Facility (GBIF: http://www.gbif.org).

Overcoming these gaps in biodiversity information has proved a serious challenge for ecologists and conservationists, while the high context-dependency of ecology makes information gaps a critical issue in conservation. Some global-level drivers are undoubtedly behind conservation problems regardless of species or space and thus should be tackled even without sufficient information (e.g., increasing food demand and climate change). It is also true, however, that local ecological phenomena are often too diverse to be predicted by general ecological theories, which need to be refined for solving specific conservation problems (Lawton 1999). Conservation practitioners usually require local- and species-level information (Braunisch et al. 2012) and inaccessibility to such relevant information can impede the use of scientific evidence in conservation (Walsh et al. 2014). Even worse, those species and countries with less information are often the more threatened in terms of conservation status (Amano and Sutherland 2013, Bland et al. 2015). Given this situation, conservation scientists have become increasingly aware of the importance of collecting and compiling scientific information that is specific to target species, locations and problems for conservation (Sutherland et al. 2004).

It is, however, still poorly known how, and whether, scientific efforts have contributed towards overcoming these information gaps over the past few decades. In this paper we focus on spatial gaps in biodiversity information, as this is one of the most studied types of information gaps in ecology and conservation (e.g., Boakes et al. 2010, Collen et al. 2008). Using data stored in the GBIF, we first quantify the geographic accumulation of biodiversity data and test how, and whether, known spatial gaps have been bridged since the 1980s. We further investigate the potential of citizen science, i.e., public involvement in research, in contributing towards overcoming these gaps. Citizen science has been suggested as an effective approach to collecting fine-grain data over continental extents as well as decadal time scales (Dickinson et al. 2010). In particular, the recent development of internet-based citizen science collecting opportunistic observation records has dramatically increased the efficiency of data collection (Sullivan et al. 2014). Most citizen science projects, however, are launched in already data-rich regions, such as North America and Europe, and thus could exaggerate existing gaps. Here we focus on one of the biggest citizen science projects collecting opportunistic observation records, eBird (Sullivan et al. 2014), as an example, and quantify its contribution to the accumulation of GBIF data.

Data accumulation has accelerated in birds, but not other taxa

The GBIF collects records on species occurrence across the globe, providing an important basis for studies in ecology and conservation (over 1400 peer-reviewed papers have been published using the GBIF data: http://www.gbif.org/mendeley). Examples include studies on the impact of climate change on species globally (Warren et al. 2013) and assessments of invasive species risk (Faulkner et al. 2014). Our earlier study indicated that the global distribution of GBIF records is similar to that of other global biodiversity databases (Amano and Sutherland 2013), thus the GBIF is a good representative for other databases. While the GBIF obviously does not store all existing occurrence records it was used as one of the largest freely accessible biodiversity databases. Though GBIF occurrence records originate from a variety of sources, including machine observation, specimen collection, fossil records and literature records, human observation accounts for over half the existing records (for more detail see http://www.gbif.org/occurrence) and, in particular, over 90% of the records collected during the last decade (Gaiji et al. 2013). We first collected the number of species occurrence records (from any sources) in each country in each year stored in the GBIF, using the occ_search function of rgbif package (Chamberlain et al. 2015) in R (R Core Team 2015) on 8th August 2015. Here each "record" represents a record of a particular species occurring in a particular country in a particular year (see examples at http://www.gbif.org/occurrence/search). Occurrence records were searched for birds (Class Aves) and other species separately.

Since 1980, the rate of increase in the number of GBIF bird records has been highest in the two data-richest regions: the Nearctic and Western Palearctic biogeographic realms, with the rate exceeding 9% per year. This is followed by the Antarctic, Australasia, Neotropic and Afrotropic realms, with a rate exceeding 4.5% per year (Figure 1a). The Eastern Palearctic, Oceania and Indo-Malay realms showed the slowest increase of below 3.0% per year. The rate of increase in the number of GBIF non-bird records was generally lower, with less variation among realms, than that of bird records (Figure 1b). Notably, the Nearctic realm showed an increase of only 1.8% per year, in contrast to 13.5% per year in bird records, while the Eastern Palearctic and Oceania realms showed a similar, or even higher, rate of increase in non-bird records compared to that in bird records (Figure 1b).

The number of GBIF bird records collected in each decade has increased over the past three decades in most realms, with a few exceptions (Figure 2a). As one example, the number of bird records collected in the Afrotropic and Antarctic realms peaked in the 1980s and 1990s, respectively (Figures 1a and 2a), mainly due to contributions from a large citizen science project covering six countries in Southern Africa (The Southern African Bird Atlas Project: Harrison et al. 1997) and two professional datasets in Antarctica, Seabirds of the Southern and South Indian Ocean (http://www.gbif.org/dataset/82dd797a-f762-11e1-a439-00145eb45e9a) and ARGOS Satellite Tracking of animals (http://www.gbif.org/dataset/82e6a41e-f762-11e1-a439-00145eb45e9a), both published by the Australian Antarctic Data Centre. In contrast, the number of GBIF non-bird records collected in each decade did not show a consistent increase in all but the Western Palearctic realms (Figure 2b). Notably, the number of non-bird records collected in the two biodiversity-rich regions, the Afrotropic and Neotropic realms, declined dramatically in the last decade compared to the preceding two decades (Figure 2b). Although stochasticity and the time lag between data collection and storage may partly explain the declines, these results suggest that scientific efforts to collect and share species occurrence data have at best not improved, and even potentially declined, in some data-poor regions despite spatial information gaps being recognised as a challenge since the 1980s (Western et al. 1989).

We also tested the relationship between the number of GBIF bird records per km^2 per species collected before (x) and during each decade (y) in each country (n = 228 countries), by fitting a power-law relationship ($y = a{\cdot}x^b$). When the exponent b exceeds 1 it signifies that the rates of record increase ($a{\cdot}x^{b-1}$) are high in originally data-rich countries while an exponent smaller than 1 is a sign that the rates are higher in data-poor countries. Although the estimated exponent did not significantly differ from 1 in the 1980s (estimate = 1.067, 95% confidence interval: 0.972 - 1.162), 1990s (1.081, 0.995 - 1.167), 2000s (0.963, 0.884 - 1.041) and 2010s (0.951, 0.878 - 1.024), the slopes for the 2000s and 2010s seem to be slightly shallower than those for the preceding two decades (Figure 3). This indicates a possibility that the rates of GBIF bird record accumulation have increased particularly in data-poor countries over the past 15 years.

Contribution from a citizen science project, eBird

eBird, launched in 2002, collects data on bird occurrence and abundance and makes the collected data available through its own platform and other biodiversity initiatives, including the GBIF (Sullivan et al. 2014). eBird might not necessarily be representative of average citizen science efforts, and of course many other citizen science projects have also submitted records to the GBIF, but it was used here to assess the contribution of one of the biggest existing citizen science projects. The number of GBIF records submitted via eBird was also collected using the occ_search function in R.

Over the past three decades, but most prominently during the last decade, eBird alone has accounted for a considerable portion of the increase in GBIF bird records, not only in the Nearctic realm, but also in the Neotropic, Indo-Malay, Eastern Palearctic and Oceania realms (shaded areas between solid and broken lines in Figure 1a and pale bars in Figure 2a). The accumulation of GBIF bird records was remarkably slow in the four (the Neotropic, Indo-Malay, Eastern Palearctic and Oceania) realms when excluding eBird contributions; the rate of increase being 1.1% per year or lower (Figure 1a). This does not mean that all records submitted via eBird would not have existed without eBird, as eBird, by providing birdwatchers with a way of keeping track of their observations and comparing their observations with others, incentivizes participants to share both new and existing data through its platform (Wood et al. 2011; as reflected in the fact that eBird has even submitted records for the 1980s and 1990s, before it was launched, Figures 1a and 2a). Nevertheless, it is certainly true that eBird has substantially increased the amount of bird occurrence data that are readily available to anyone in the world.

Although the increase in eBird records after 2010 and the number of GBIF bird records available in 2009 also showed a power-law relationship with an exponent of 1 (0.950, 0.869 – 1.031), eBird records account for more than a half of all GBIF bird records recorded after 2010 in 16 of the 20 data-poorest countries (highlighted in blue in Figure 4, see Table 1 for details). These results highlight the potential of citizen science to aid data collection and sharing even in data-poor regions.

Potentials and limitations of citizen science in tackling information gaps

This study derived two important findings: (i) the accumulation of GBIF bird records has accelerated dramatically over the past three decades, even in some data-poor regions, which is at least partly attributable to contributions from eBird, and (ii) the rate of increase in GBIF non-bird records is generally low compared to bird records, and is even slowing in some data-poor regions, such as the Afrotropic and Neotropic realms.

For birds, the proportion of eBird-derived data in GBIF records was surprisingly high across the globe (Figure 4), which is likely to be attributable to contributions from both (i) the recent growth of local birding communities, notably in the Neotropic and Indo-Malay realms, but also in other regions, and (ii) birders in more economically-developed countries (e.g., the US) visiting other countries. There are other similar programs that collect opportunistic observation records in different regions, and thus have the potential to further aid data collection and compilation globally. In Europe, the Euro Bird Portal (http://www.eurobirdportal.org/ebp/en/) has been established to represent a European data repository based on aggregated data from 13 online bird recording portals from across Europe, collecting about 30 million bird records every year. In Africa, the Second Southern African Bird Atlas Project is currently underway and has already collected about seven million occurrence records, by taking advantage of specifically-developed mobile apps (http://sabap2.adu.org.za/). In Australia, the Eremaea Birds (http://www.eremaea.com/), which merged with eBird in 2014, has submitted about 2.7 million records to the GBIF (http://www.gbif.org/publisher/633f217c-c007-48dc-86ed-f8fdae6fd0d8). WikiAves (http://en.wikiaves.com/) has also collected 1.5 million bird occurrence records (as photos submitted from users) in Brazil. So, while this study was focused on eBird, the contribution of citizen science projects as a whole to the global accumulation of species occurrence data can be even higher. To ensure increased accessibility, these data are also encouraged to be shared through the GBIF.

Given the slow rate of increase in non-bird records in most parts of the world, the question is whether citizen science projects are similarly effective in tackling spatial information gaps in other taxa. Citizen science is generally biased towards vertebrates and terrestrial ecosystems (Theobald et al. 2015) and even within each taxon, towards particular species groups (e.g., easily observable species). We need an assessment of the taxa and species that merit more attention and where data accumulation should be encouraged. Newer citizen science projects may serve a similar function as eBird for other species in the near future. For example, iNaturalist (http://www.inaturalist.org/), whose contribution to the GBIF has been increasing exponentially, has already been used specifically for collecting occurrence records of amphibians (http://www.inaturalist.org/projects/global-amphibian-bioblitz) and freshwater fish (http://www.inaturalist.org/projects/global-freshwater-fish-bioblitz) globally. Some projects, such as iNaturalist and iSpot (http://www.ispotnature.org/), do not necessarily require each observer to have identification skills, as they “crowdsource” species identification, and thus have a potential to produce high volumes of data for diverse taxa (Pimm et al. 2014). eBird has successfully incentivized participants by providing tools to keep track, view and compare their observations (Wood et al. 2011), which, if incorporated, may also encourage the collection of data on other taxa. Indirect ways of observations, such as acoustic monitoring (Walters et al. 2012), camera traps (Ahumada et al. 2011), the use of environmental DNA (Thomsen and Willerslev 2015) and photos posted on social networking sites (Barve 2014), can also be powerful approaches for otherwise less detectable species. One serious challenge though is the lack of common, comprehensive taxonomy in most organisms other than a few charismatic groups, such as birds, mammals and some higher plants (Secretariat of the Convention on Biological Diversity 2007). This reemphasises the importance of taxonomy and reviving it for biodiversity conservation (Pearson et al. 2011).

It is also true, however, that many citizen science projects are opportunistic and not specifically aimed at bridging spatial information gaps. Even eBird has collected few records in some of the data-poorest countries (Table 1). To tackle spatial information gaps more effectively, it is therefore crucial to better understand the causes of data scarcity in some regions. Potential factors that have been suggested to cause spatial information gaps include wealth, insufficient expertise, infrastructure and communication, and inaccessibility due to geographical location and/or security level (Amano and Sutherland 2013, Collen et al. 2008, Martin et al. 2012). The level of concern and attitudes for environmental issues (Franzen and Vogl 2013), while having attracted less attention, may be another important driver of ecological data collection. These factors can also cause an unequal distribution of data even within data-rich countries (Isaac and Pocock 2015). Understanding these barriers to the global collection and compilation of biodiversity data can help us overcome some of the barriers (see for example Extreme Citizen Science: https://www.ucl.ac.uk/excites), or at least incorporate knowledge of identified constraints as well as current data coverage in introducing spatial prioritisation to future efforts of data collection.

Meanwhile, not all existing data are effectively shared at the global level. One factor may simply be the lack of sufficient communication. Given that both language and geographical locations of program hosts can represent barriers to global data compilation (Amano and Sutherland 2013), it should be effective to establish partnerships with local projects in data-poor regions (as eBird does), develop new projects using local languages and translate existing programs into different languages. Such multi-lingualisation with user-friendly online platforms would also help gain access to historical data in data-poor regions. The lack of an open access culture in some regions is another barrier to the global compilation of existing data (Hobern et al. 2013). For example, large volumes of occurrence observations have been accumulated in Japan for a range of taxa since the 1970s, but many such data have not been shared on the GBIF yet, possibly due to several issues including the absence of a data-sharing culture (Osawa et al. 2014). A priority in such a situation will be to foster a culture of data sharing, for example, through public funding and other incentives, along with the proper recognition of its advantages.

We also need to be aware of the limitations of citizen science data, notably the types and quality of currently available data. For example, the collection of long-term abundance data in data-poor regions to date has largely been limited to birds (e.g., International Waterbird Census: http://www.wetlands.org/OurWork/Biodiversity/Monitoringwaterbirdpopulations/tabid/773/Default.aspx). One clear challenge is to evaluate, whether through novel surveys or with modelling, changes in species abundance over space and time, which are a central focus in biodiversity conservation. For information other than species occurrence and abundance, iNaturalist has already been used to collect data on species interactions (Poelen et al. 2014) and behaviour (Sheehan et al. 2015). Assuring data quality in citizen science projects requires careful design of data input and management procedures (Sullivan et al. 2014), training of surveyors and standardization of methods (Mackechnie et al. 2011). Recording associated information, such as sampling effort (Pimm et al. 2014), identification uncertainties, species absence (Sullivan et al. 2014), species interactions (Dickinson et al. 2010) and environmental and social information (Crain et al. 2014), would also improve the usability of data and reliability of inferences derived. To this end, the GBIF is starting to store "sample-based" data, which include information on the quantity of organisms and sampling efforts (http://www.gbif.org/page/82105).

Finally, while conservation practices usually require local- and species-level information, the amount of effort and time that can be spared for data collection is inevitably limited. We thus call for the need to conduct thorough discussions on which areas, taxa and data types should be prioritised for future efforts of data collection. At the practical level, prioritising the collection of data that are truly needed for conservation requires feedback from data users. Global initiatives, such as the Intergovernmental Platform on Biodiversity and Ecosystem

Services (http://www.ipbes.net/), provide an ideal opportunity for this purpose, and could not just simply identify information gaps as a problem but could also actively engage in solving it, for example, by scanning the types of data required and encouraging projects—run both by citizens and professionals— that collect those priority data types. It would also be effective to pursue how we can complement the lack of information with modelling approaches. Some attempts, such as testing model transferability over space (e.g., Randin et al. 2006) and informing the conservation of data-deficient species with predictive modelling (e.g., Bland et al. 2015), have already been made and should be encouraged further.

Acknowledgments

T.A. was supported by the European Commission's Marie Curie International Incoming Fellowship Programme (PIIF-GA-2011-303221) and the Isaac Newton Trust and W.J.S. by the Arcadia Fund. Thanks to Timothy Beardsley, René van der Wal and two anonymous reviewers for their comments on an earlier draft and M. Amano for all the support.

References cited

Ahumada JA, et al. 2011. Community structure and diversity of tropical forest mammals: data from a global camera trap network. Philosophical Transactions of the Royal Society of London B: Biological Sciences 366: 2703-2711.

Amano T, Sutherland WJ. 2013. Four barriers to the global understanding of biodiversity conservation: wealth, language, geographical location and security. Proceedings of the Royal Society B: Biological Sciences 280: 20122649.

Barve V. 2014. Discovering and developing primary biodiversity data from social networking sites: A novel approach. Ecological Informatics 24: 194-199.

Bland LM, Collen BEN, Orme CDL, Bielby JON. 2015. Predicting the conservation status of data-deficient species. Conservation Biology 29: 250-259.

Boakes EH, McGowan PJK, Fuller RA, Ding C-q, Clark NE, O'Connor K, Mace GM. 2010. Distorted views of biodiversity: spatial and temporal bias in species occurrence data. PLoS Biology 8: e1000385.

Braunisch V, Home R, Pellet J, Arlettaz R. 2012. Conservation science relevant to action: a research agenda identified and prioritized by practitioners. Biological Conservation 153: 201-210.

Chamberlain S, Ram K, Barve V, Mcglinn D. 2015. rgbif: Interface to the Global Biodiversity Information Facility API. R package version 0.8.8.: http://CRAN.R-project.org/package=rgbif.

Collen B, Ram M, Zamin T, McRae L. 2008. The tropical biodiversity data gap: addressing disparity in global monitoring. Tropical Conservation Science 1: 75-88.

Crain R, Cooper C, Dickinson JL. 2014. Citizen science: a tool for integrating studies of human and natural systems. Annual Review of Environment and Resources 39: 641-665.

Dickinson JL, Zuckerberg B, Bonter DN. 2010. Citizen science as an ecological research tool:

challenges and benefits. Annual Review of Ecology, Evolution, and Systematics 41: 149-172.

Faulkner KT, Robertson MP, Rouget M, Wilson JRU. 2014. A simple, rapid methodology for developing invasive species watch lists. Biological Conservation 179: 25-32.

Franzen A, Vogl D. 2013. Two decades of measuring environmental attitudes: a comparative analysis of 33 countries. Global Environmental Change 23: 1001-1008.

Gaiji S, Chavan V, Ariño AH, Otegui J, Hobern D, Sood R, Robles E. 2013. Content assessment of the primary biodiversity data published through GBIF network: status, challenges and potentials. Biodiversity Informatics 8: 94-172.

Gardner JL, Amano T, Sutherland WJ, Joseph L, Peters A. 2014. Are natural history collections coming to an end as time-series? Frontiers in Ecology and the Environment 12: 436-438.

Harrison JA, Allan DG, Underhill LG, Herremans M, Tree AJ, Parker V, Brown CJ, eds. 1997. The atlas of southern African birds. Vol. 1: Non-passerines Johannesburg: BirdLife South Africa.

Hobern D, et al. 2013. Global biodiversity informatics outlook: delivering biodiversity knowledge in the information age: Global Biodiversity Information Facility (Secretariat).

Isaac NJB, Pocock MJO. 2015. Bias and information in biological records. Biological Journal of the Linnean Society 115: 522-531.

Isaac NJB, van Strien AJ, August TA, de Zeeuw MP, Roy DB. 2014. Statistics for citizen science: extracting signals of change from noisy ecological data. Methods in Ecology and Evolution 5: 1052-1060.

IUCN. 2015. The IUCN Red List of Threatened Species. Version 2015.2: http://www.iucnredlist.org. Downloaded on 06 August 2015.

Lawton JH. 1999. Are there general laws in ecology? Oikos 84: 177-192.

Mackechnie C, Maskell L, Norton L, Roy D. 2011. The role of 'Big Society' in monitoring the state of the natural environment. Journal of Environmental Monitoring 13: 2687-2691.

Martin LJ, Blossey B, Ellis E. 2012. Mapping where ecologists work: biases in the global distribution of terrestrial ecological observations. Frontiers in Ecology and the Environment 10: 195-201.

Osawa T, Jinbo U, Iwasaki N. 2014. Current status and future perspective on "Open Data" in biodiversity science, Japan. Japanese Journal of Ecology 64: 153-162 (in Japanese).

Pearson DL, Hamilton AL, Erwin TL. 2011. Recovery plan for the endangered taxonomy profession. BioScience 61: 58-63.

Pimm SL, Jenkins CN, Abell R, Brooks TM, Gittleman JL, Joppa LN, Raven PH, Roberts CM, Sexton JO. 2014. The biodiversity of species and their rates of extinction, distribution, and protection. Science 344: 1246752.

Poelen JH, Simons JD, Mungall CJ. 2014. Global biotic interactions: An open infrastructure to share and analyze species-interaction datasets. Ecological Informatics 24: 148-159.

R Core Team. 2015. R: A language and environment for statistical computing. Vienna, Austria (http://www.R-project.org/): R Foundation for Statistical Computing.

Randin CF, Dirnböck T, Dullinger S, Zimmermann NE, Zappa M, Guisan A. 2006. Are niche-based species distribution models transferable in space? Journal of Biogeography 33: 1689-1703.

Secretariat of the Convention on Biological Diversity. 2007. Guide to the Global Taxonomy Initiative. Technical Series no. 30. Montreal: Secretariat of the Convention on Biological Diversity.

Sheehan MJ, Botero CA, Hendry TA, Sedio BE, Jandt JM, Weiner S, Toth AL, Tibbetts EA. 2015. Different axes of environmental variation explain the presence vs. extent of cooperative nest founding associations in Polistes paper wasps. Ecology Letters: n/a-n/a.

Sullivan BL, et al. 2014. The eBird enterprise: an integrated approach to development and application of citizen science. Biological Conservation 169: 31-40.

Sutherland WJ, Pullin AS, Dolman PM, Knight TM. 2004. The need for evidence-based conservation. Trends in Ecology and Evolution 19: 305-308.

Theobald EJ, et al. 2015. Global change and local solutions: tapping the unrealized potential of citizen science for biodiversity research. Biological Conservation 181: 236-244.

Thomsen PF, Willerslev E. 2015. Environmental DNA - an emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation 183: 4-18.

Walsh JC, Dicks LV, Sutherland WJ. 2014. The effect of scientific evidence on conservation practitioners' management decisions. Conservation Biology: n/a-n/a.

Walters CL, et al. 2012. A continental-scale tool for acoustic identification of European bats. Journal of Applied Ecology 49: 1064-1074.

Warren R, et al. 2013. Quantifying the benefit of early climate change mitigation in avoiding biodiversity loss. Nature Climate Change 3: 678-682.

Webb TJ, Mindel BL. 2015. Global patterns of extinction risk in marine and non-marine systems. Current Biology 25: 506-511.

Western D, Pearl MC, Pimm SL, Walker B, Atkinson I, Woodruff DS. 1989. An agenda for conservation action. Pages 304-323 in Western D, Pearl M, eds. Conservation for the twenty-first century. New York: Oxford University Press.

Wood C, Sullivan B, Iliff M, Fink D, Kelling S. 2011. eBird: engaging birders in science and conservation. PLOS Biology 9: e1001220.

Table 1. Top 20 countries/territories (names based on the ISO 3166-1) with the fewest Global Biodiversity Information Facility (GBIF) bird records per km^2 per species in 2009. Countries/territories where eBird records account for more than a half of the increase in GBIF records since 2010 are shown in bold. Note that records taken in 2014 and 2015 were not included here.

Country	Biogeographic Realm	Area (km^2)	Bird species richness	Number of GBIF bird records recorded before 2010 (km^{-2} species^{-1})	Number of GBIF bird records recorded after 2010	Number of eBird records recorded after 2010 (% of all GBIF records)
Libya	**W Palearctic**	**1759540**	**330**	**239 (4.12×10^{-7})**	**1242**	**1241 (100%)**
Mali	Afrotropic	1240192	618	1092 (1.42×10^{-6})	17	0 (0%)
Chad	**Afrotropic**	**1284000**	**548**	**1052 (1.50×10^{-6})**	**145**	**144 (99%)**
Niger	**Afrotropic**	**1267000**	**502**	**961 (1.51×10^{-6})**	**32**	**16 (50%)**
Russia	**E Palearctic**	**17098242**	**1425**	**40908 (1.68×10^{-6})**	**9256**	**6767 (73%)**
Mauritania	W Palearctic	1030700	529	921 (1.69×10^{-6})	967	90 (9%)
Turkmenistan	**E Palearctic**	**488100**	**382**	**341 (1.83×10^{-6})**	**10**	**10 (100%)**
Tajikistan	**E Palearctic**	**143100**	**358**	**94 (1.83×10^{-6})**	**30**	**29 (97%)**
Burkina Faso	**Afrotropic**	**274200**	**477**	**391 (2.99×10^{-6})**	**189**	**110 (58%)**
Western Sahara	W Palearctic	266000	195	195 (3.76×10^{-6})	2823	1 (0%)
Belarus	W Palearctic	207600	312	247 (3.81×10^{-6})	734	27 (4%)
Sudan	**Afrotropic**	**1861484**	**961**	**7693 (4.30×10^{-6})**	**831**	**767 (92%)**
Kazakhstan	**E Palearctic**	**2724900**	**497**	**6363 (4.70×10^{-6})**	**4715**	**4707 (100%)**
Central African Republic	**Afrotropic**	**622984**	**725**	**2374 (5.26×10^{-6})**	**273**	**273 (100%)**
Algeria	**W Palearctic**	**2381741**	**384**	**5849 (6.40×10^{-6})**	**83**	**80 (96%)**
China	**E Palearctic**	**9596961**	**1273**	**79035 (6.47×10^{-6})**	**26653**	**25214 (95%)**
Yemen	**W Palearctic**	**527968**	**419**	**1516 (6.85×10^{-6})**	**32**	**31 (97%)**
Nigeria	**Afrotropic**	**923768**	**909**	**6595 (7.85×10^{-6})**	**4339**	**4241 (98%)**
Benin	**Afrotropic**	**112622**	**539**	**546 (8.99×10^{-6})**	**3073**	**2608 (85%)**
Uzbekistan	**E Palearctic**	**447400**	**365**	**1481 (9.07×10^{-6})**	**127**	**86 (68%)**

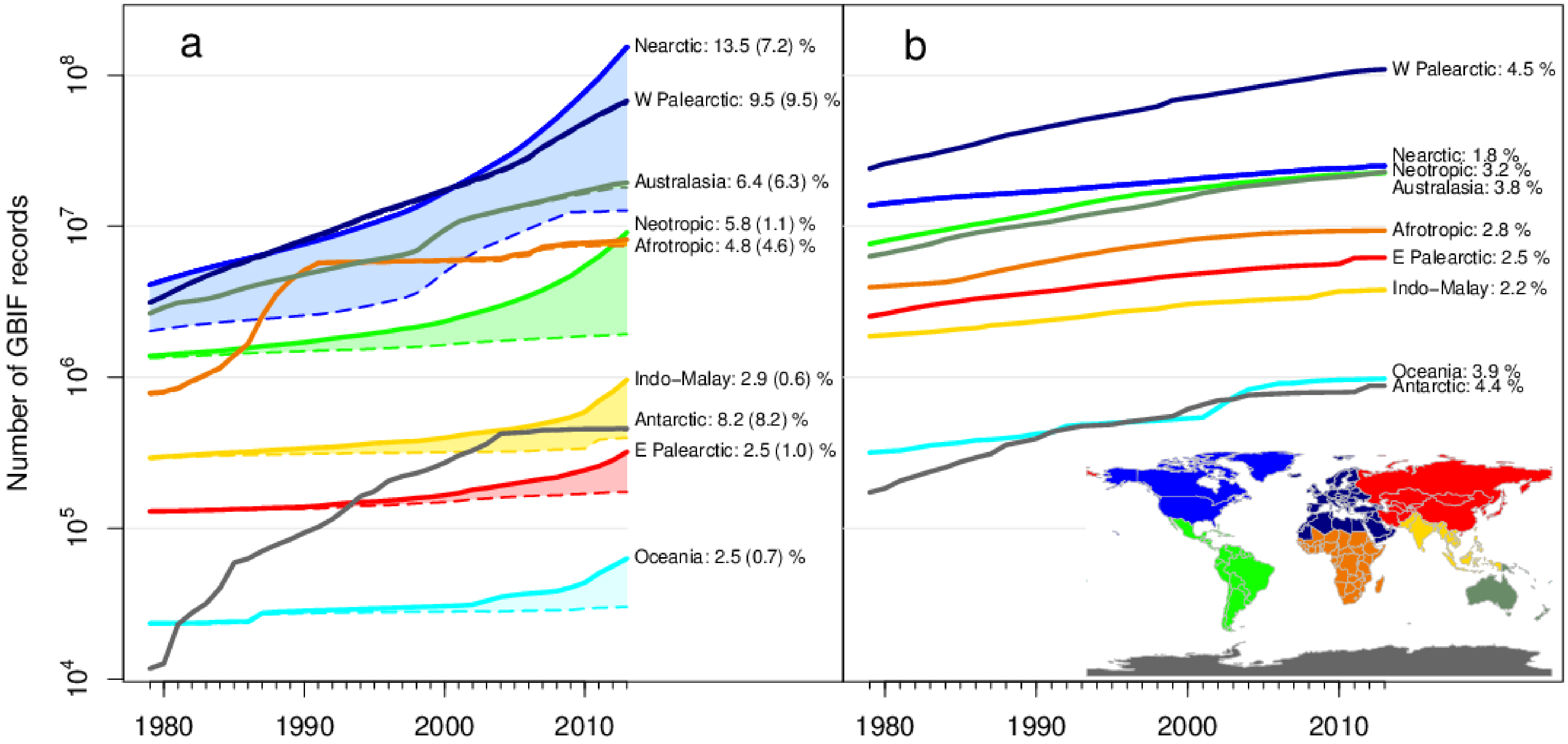

Figure 1. Changes in the cumulative number of occurrence records of (a) birds and (b) non-bird species in the Global Biodiversity Information Facility between 1979 and 2013. Solid lines indicate the total number of records by biogeographic realms. In (a), broken lines represent the number of records without contributions from eBird, with shaded areas showing the number of records submitted via eBird. Note that broken lines are almost invisible for the Western Palearctic and Antarctic realms, as the contribution of eBird there was small. The values next to the names of biogeographic realms represent the annual growth rate of the number of bird records with and without (in parentheses) eBird in (a) and the annual growth rate of the number of non-bird records in (b) over the past 34 years. Note that the y-axis is on a log-scale.

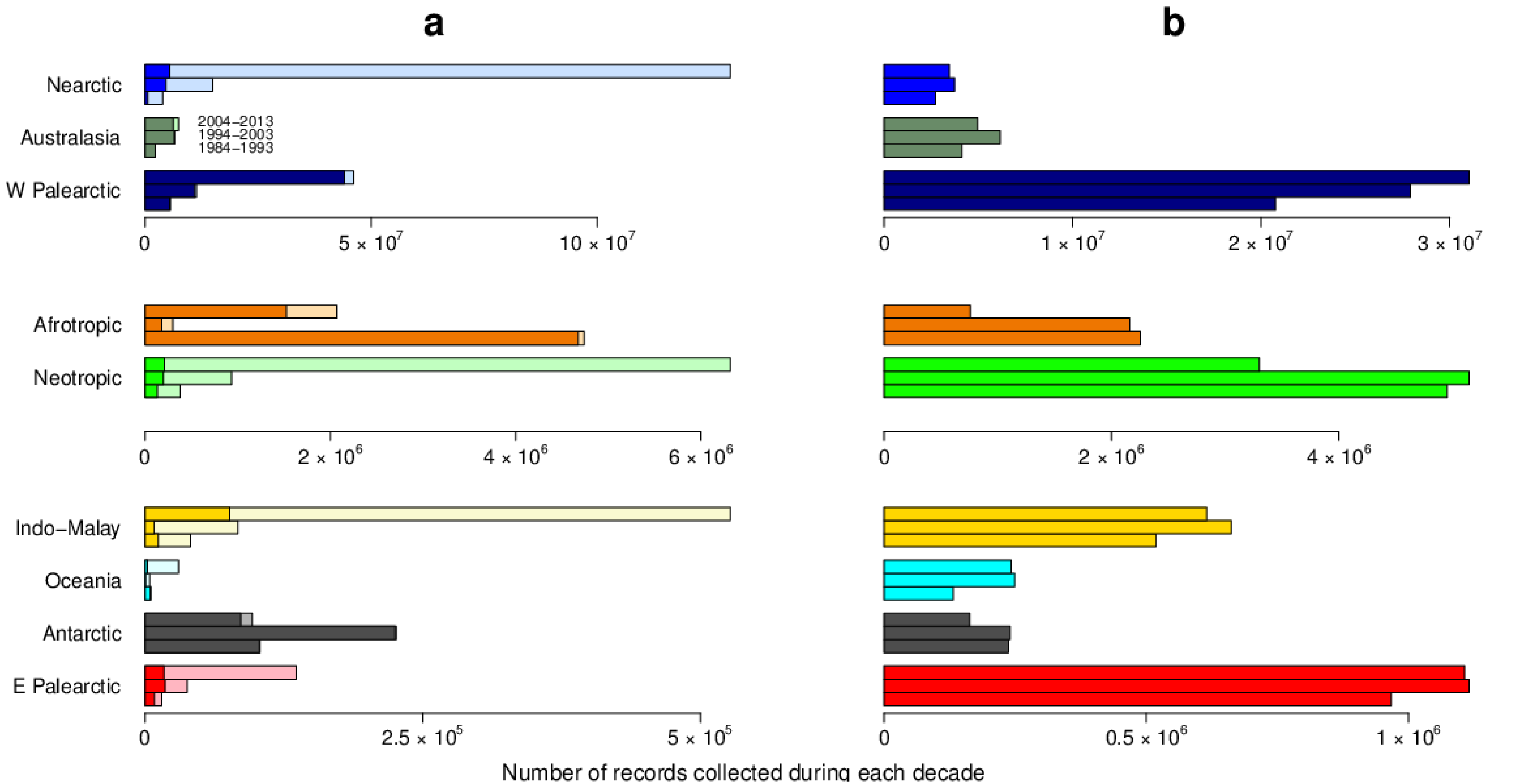

Figure 2. The number of occurrence records of (a) birds and (b) non-bird species in the Global Biodiversity Information Facility, collected during each decade (in each biogeographic realm from the bottom, 1984-1993, 1994-2003, and 2004-2013). In (a) parts shown in pale colours indicate the number of records submitted via eBird.

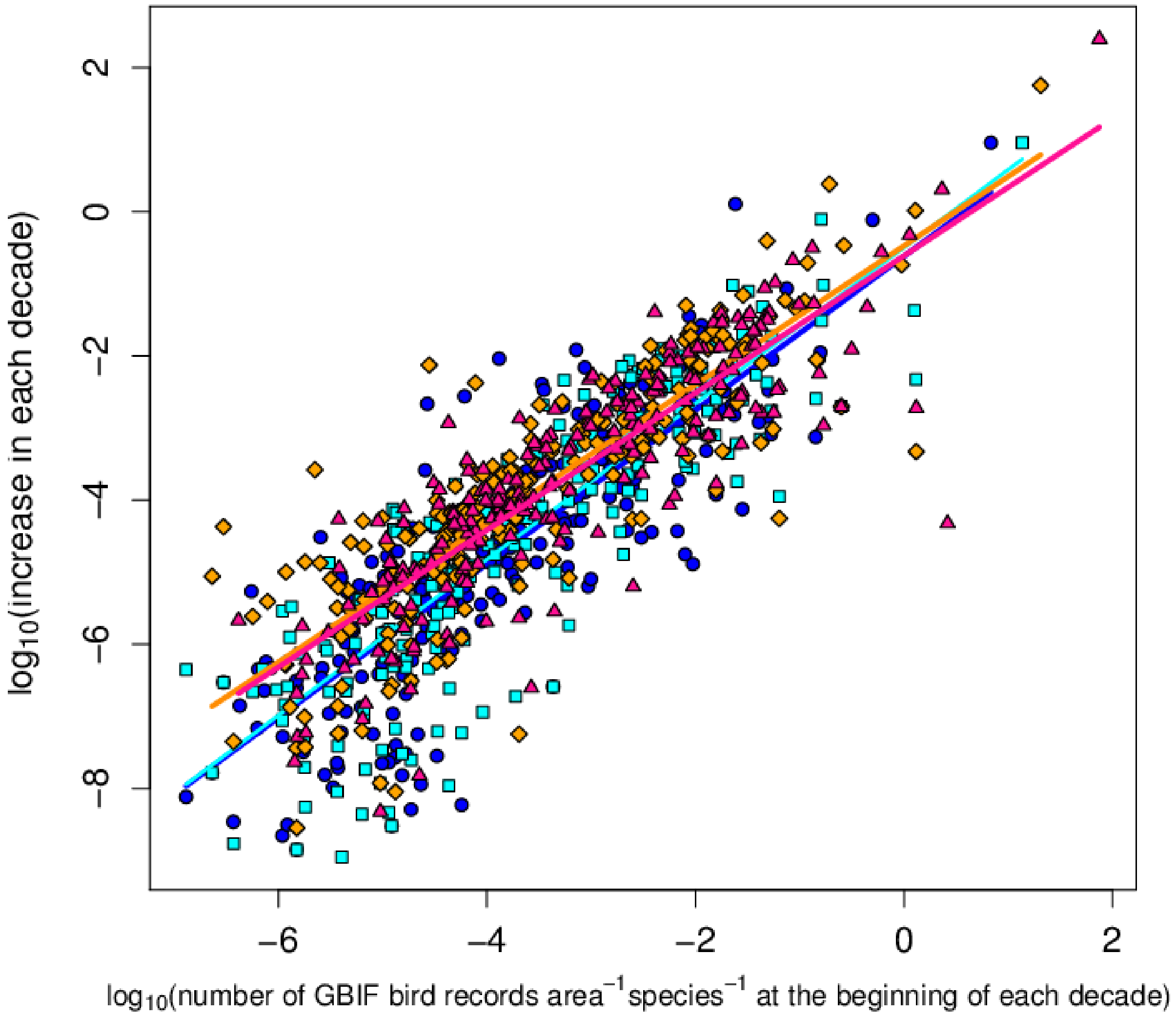

Figure 3. The relationship between the number of the Global Biodiversity Information Facility (GBIF) bird records per area per species at the beginning of each decade and increase in that decade by country (1980s: circles in blue, 1990s: squares in sky blue, 2000s: diamonds in orange, 2010s: triangles in pink). Regression lines are also shown (see main text for the estimated slopes).

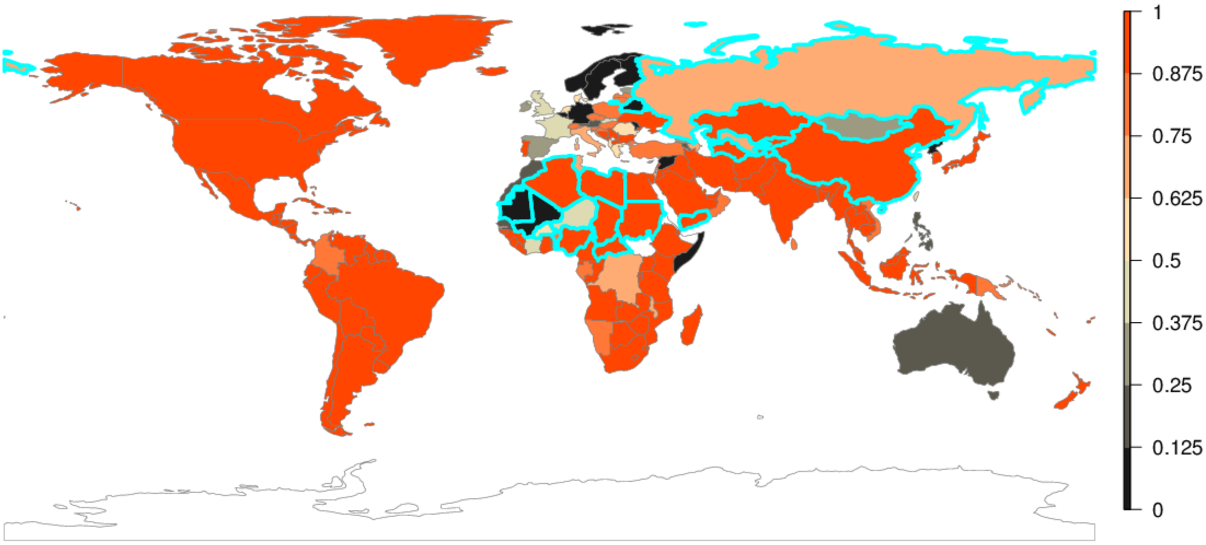

Figure 4. The proportion of eBird records in the increase in the Global Biodiversity Information Facility (GBIF) bird records between 2010 and 2013. Contribution from eBird is particularly high in countries shown in orange. Twenty countries with fewest GBIF records per km^2 per species in 2009 outlined blue.